Abel Inversion Using Total Variation Regularization: Applications

LA-UR-05-2657

April 14, 2005

Abstract

We apply total-variation (TV) regularization methods to Abel inversion tomography. Inversions are performed using the fixed-point iteration method and the regularization parameter is chosen such that the resulting data fidelity approximates the known or estimated statistical character of the noisy data. Five one dimensional examples illustrate the favorable characteristics of TV regularized solutions: noise inversion suppression and density discontinuity preservation. One experimental and two simulated examples from X-ray radiography also illustrate limitations due to a linear projection approximation. TV regularized inversions are shown to be superior to squared gradient regularized inversions for objects with density discontinuities. We also introduce an adaptive TV method that utilizes a modified discrete gradient operator acting only apart from data-determined density discontinuities. This method provides improved density level preservation relative to the basic TV method.

Keywords: Abel Transform, Inverse Problems, Total Variation, Tomography, Regularization.

AMS 2000 MSCC: (pending)

1 Introduction

The density reconstruction of objects from several radiogarphic views is a classic and important tomography problem. The inversion is typically formulated using the Radon Transform or related approaches [3] and is regularized using any one of a number of techniques. For a good introduction to regularization of inverse problems see Vogel[4]. For objects with cylindrical symmetry the inversion reduces to the inverse Abel transform. In this paper we consider objects with one dimensional intrinsic property descriptions $\rho(r)$. We leave discussions of applications to objects of general cylindrical symmetry $\rho(r, z)$ to a future paper. The continuous Abel transform is

$$d(x) = 2\int_{|x|}^{\infty} \frac{r\rho(r)}{\sqrt{r^2 - x^2}}\, dr, \tag{1}$$

where d is a line-integral density relative to ρ. For example, if ρ is a volumetric density then d is an areal density. Equation (1) has a well-defined inverse: if d is given by Equation (1), then

$$\rho(r) = -\frac{1}{\pi r}\frac{d}{dr}\int_{r}^{\infty} \frac{xd(x)}{\sqrt{x^2 - r^2}}\, dx. \tag{2}$$

A discretized version of Equation (2) is the basis for Abel inversion tomography.

In practice there are a number of difficulties to address. First, Equation (1) is a simplified description of typically very nonlinear experiments. Second, the inverse problem can be ill-posed. Since radiographs are transmission intensity maps (or some equivalent), the corresponding intrinsic material property is an attentuation coefficient. Thus, obtaining an object density requires an additional transformation either in the radiographic space (intensity to areal density) or in the object description (attenuation to density). Both approaches are nontrivial requiring a detailed understanding of the physics of interaction

between the object and detectors with the radiation. The simplest case, reconstructing an object of a single material, can usually be solved with good accuracy. However, multiple-material situations cause this additional transformation to be non-unique, and significant prior information about the object is necessary before the inverse Abel transform can be utilized for quantitative evaluation.

A third difficulty is the ill-conditioning of the discrete inverse Abel transform. While a given inversion is unique, small perturbations in d lead to large deviations in ρ. This is because Equation (2) defines an unbounded operator (see [1] for details). The inversion must be regularized to obtain meaningful results from noisy data.

The discretized Abel transform can be formulated as a matrix P. If we consider the object radial density values as a vector ρ of n elements and the areal density projections as a vector d of m elements, then P is a non-sparse $m \times n$ matrix. This projection is invertible if $n \leq m$, but is poorly conditioned when $n \sim m$.

The discrete regularized inverse Abel transform is formulated as a functional minimization problem:

$$\min_{\rho} F(\rho) = \min_{\rho} \left\{ ||P\rho - d||_{\mathrm{df}} + \alpha R(\rho) \right\}, \tag{3}$$

where $|| \cdot ||_{\mathrm{df}}$ is an appropriate data fidelity norm, and $R(\rho)$ is a regularization term determined by a probability model of the types of objects we expect.

In this paper we focus on the use of regularization to address the ill-conditioned inverse problem. We demonstrate, by example, the importance of the choice of R for recovering object density profiles ρ from noisy areal density data d.

2 Methods

We compute P as the parallel planar projection from a two-dimensional object space onto a one-dimensional data space. Noise is treated as stationary Gaussian white noise. In practice this approach works well, but typical experiments are dominated by locally-Poisson noise statistics. The examples that follow have different noise characteristics, so our treatment of data fidelity can be expected to have mixed results. We consider functionals of the form

$$F(\rho) = \frac{1}{2}\int_0^\infty dr\,|P\rho - d|^2 + 2\pi\alpha \int_0^\infty r\,dr\,|\nabla\rho|^p. \tag{4}$$

In particular we consider $p \in \{0, 1, 2\}$, corresponding to regularization types none, total variation (TV), and H^1 respectively. We also introduce a modified TV regularization method that utilizes an adaptive gradient operator. The minimizing solution of Equation (3) depends on the choice of α as well as the regularizer $R(\rho)$. The α we select is that which leads to a solution with data fidelity norm equal to the known or estimated variance in the data noise. Many problems may also benefit by a more careful approach to data fidelity modelling that assumes correct statistics. For example, it is clear that most radiography applications are governed by Poisson statistics. Such treatments are outside the scope of this paper.

Our approach is to use the largest invertible discrete linear projection and regularize the inversion. The data resolution determines the size of the projection operator: the discretization of P is an $m \times m$ matrix and the object is reconstructed at the same resolution. The methods outlined in the following sections are not limited to this invertabillity condition.

2.1 Inversion without regularization

The unregularized inversion is the $p = 0$ case of Equation (4). This lack of regularization is appropriate for situations in which no prior object knowledge is available. The object density reconstruction is given explicitly by the psuedo-inverse $\rho = (P^T P)^{-1} P^T d$. It is expected to produce poor results in real scenarios due to the combination of noisy data and the ill-conditioning. In particular, we expect noise in the most stable directions of P (the most unstable of P^{-1}) to be amplified significantly and lead to useless inversions.

2.2 H^1 regularization

The $p = 2$ case of Equation (4) is the H^1-regularization minimization. It can be shown to be equivalent to applying the diffusion operator on the unregularized solution. The equivalent diffusion time is inversely related to α. The result is a smooth reconstruction clearly biased against discontinuities in ρ. The same is true for all $p > 1$ solutions. The object density reconstruction is given explicitly by

$$\rho = \left(P^T P - 4\pi\alpha r\nabla^2 - 4\pi\alpha\nabla\right)^{-1} P^T d. \tag{5}$$

For the gradient operator we use a simple forward differencing with Neumann boundary conditions. If the inverse is nearly singular it is advantagous to use the solution methods outlined in the next subsection.

2.3 TV regularization

The $p = 1$ case of Equation (4) is the TV regularization minimization. This regularization is not biased against density discontinuities; it penalizes such density edges by a reduction in amplitude but not at the expense of smoothing

the edge. Following Vogel [4], we use the lagged-diffusivity fixed point method to find the minimum of $F(\rho)$. We compute the gradient and an approximate Hessian of the discretized version of the cost functional F that define a quasi-Newton step towards the minimum. From equation (4),

$$F'(\rho) = P^T(P\rho - d) + \alpha L(\rho)\rho, \tag{6}$$

and

$$F''(\rho) = P^T P + \alpha L(\rho) + \alpha L'(\rho)\rho, \tag{7}$$

where we adopt the notation $R'(\rho) = L(\rho)\rho$. In particular, if $p = 2$ then $L(\rho) = 2\pi\left(r\nabla^2 + \frac{\partial}{\partial r}\right)$. If $p = 1$ the non-differentiability of R when $\nabla\rho = 0$ is problematic. The difficulty is overcome by choosing a suitably small parameter β and defining R by

$$R(\rho) = 2\pi\int_0^\infty r dr\sqrt{|\nabla\rho|^2 + \beta^2}. \tag{8}$$

Now R has a well-defined derivative from which $L(\rho)$ is identified:

$$L(\rho) = 2\pi\left[r\nabla\cdot\left(\frac{\nabla}{\sqrt{|\nabla\rho|^2+\beta^2}}\right) + \left(\frac{\frac{\partial}{\partial r}}{\sqrt{|\nabla\rho|^2+\beta^2}}\right)\right]. \tag{9}$$

Since $L(\rho)$ depends explicitly on the solution ρ, we let the iterate solution $\rho_{\nu+1}$ depend on the previous solution ρ_ν. This is the origin of the term lagged diffusivity. The quasi-Newton iteration is given by

$$\begin{aligned}\rho_{\nu+1} &= \rho_\nu - [\mathrm{Hess}\, F(\rho)]^{-1}\nabla F(\rho_\nu)\\ &\approx \rho_\nu - \left[P^T P + \alpha L(\rho_\nu)\right]^{-1}\left[P^T(P\rho_\nu - d) + \alpha L(\rho_\nu)\rho_\nu\right].\end{aligned} \tag{10}$$

where we use the approximate Hessian which omits the $L'(\rho)\rho$ term. While Newton methods are expected to have quadratic convergence, the use of a lagged-diffusivity and approximate Hessian guarantee only linear convergence [2]. Methods with faster convergence (for example, primal dual) require larger computational overhead. For the small problems of interest here ($n < 10^3$), linear convergence is sufficient.

The discretization of Equation (10) is straightforward. First, we define the diffusivity function

$$\phi = \frac{1}{\sqrt{|\nabla\rho|^2 + \beta^2}}, \tag{11}$$

so that

$$L(\rho) = \phi\frac{\partial}{\partial r} + r\nabla\cdot\phi\nabla. \tag{12}$$

We then use simple forward differencing to construct the discrete gradient operator. The use of second order methods can present difficulties when attempting to recover discontinuities. See [1] for a detailed discussion of the functional F when $p = 1$.

2.4 Adaptive TV regularization

Since most static objects have only piecewise continuous densities, we would like a regularizer with the edge-preserving quality of TV and absolute amplitude smoothing qualities of H^1. To this end, we present a fourth regularization method which we designate as adaptive total variation (aTV). In this method we mask the discrete gradient operator along an edge location set that identifies object discontinuities. This edge location set is determined by dynamically adjusting the regularization parameter α and a discontinuity threshhold. Both

are initially large and are gradually reduced over several iterations. A final H^1 regularization is performed with the fully-masked gradient operator. Thus, regions of constant or smoothly varying density are smoothed and edge locations and amplitudes are preserved.

3 Examples and Discussion

We show results of the four regularized inversion methods applied to both synthetic and real data. In all cases we work with objects of one-dimensional description ρ and corresponding one-dimensional data d.

3.1 Synthetic Example 1

The first example is the reconstruction of an object of ten nested varying density rings. The object, data, and reconstructions are shown in Figure 1. The object density profile (a) is given by $n = 200$ ring densities. The corresponding projection data (b) is a synthetically generated areal density with added Gaussian noise. The variance of the noise is a uniform 1.5 percent of the maximum noiseless data value. In each of the remaining subfigures, the actual object is shown in black and a reconstruction is shown in red: (c) unregularized; (d) H^1; (e) TV; and (f) aTV. As expected, the unregularized reconstruction is overwhelmed by noise amplification through the ill-conditioned inverse projection [4]. A smoothing regularizer does very well at reducing high frequency noise artifacts and even preserves the general character of the object. But the same regularizer is unable to capture density discontinuities; smoothing the edges. The TV regularizer captures many of the discontinuities and provides a visually pleasing reconstruction. The aTV regularization procedure does the best of all. It best identifies all discontinuities and best preserves the actual density levels

to within that given by the local data.

3.2 Synthetic Example 2

The second example is the reconstruction of an object with a mixture of piecewise smooth and piecewise constant density variations. The object, data, and reconstructions are shown in Figure 2. The object is of similar description to the previous object. This object is more complex in that it has some regions of smoothly varying density and density discontinuities are not regularly spaced. The data is synthetic with added Gaussian noise of variance 1.0 percent of the maximum noiseless data value. The reconstructions are ordered as in the previous example. We note that the TV-based regularizers again perform better. Even the smallest object features are partially recovered. This suggests that our regularizer and data fidelity metric are excellent choices.

3.3 Simulated Example 1

The third examples is an edge-detection exercise from a simulated X-ray radiograph of a set of nested spherical shells. Figure 3 shows the radiograph (a), computed radial areal density (b), and reconstructions (c) and (d). Radiographs were numerically simulated using the Monte Carlo code MCNP[1] with X-ray scattering suppressed. The simulated source is a bi-chromatic cobalt-60 source set at a distance sufficient to adequately approximate a parallel beam. The object is a set of four nested spheres whose boundary locations are indicated by vertical grid lines in the reconstruction figures. In this case we must convert the radiograph from intensity (I) to areal density. Without explicit

[1]MCNP is a particle transport code that includes detailed physics treatments of photoatomic and photo-electron interactions. It correctly treats Bremsstrahlung photons created from recoil or ejected electrons as well as photon scattering. It allows high-fidelity simulations of actual experimental conditions including source particulars and experiment geometry. Our simulated radiographs were provided by Jeff Favorite of the Applied Physics Division at Los Alamos National Laboratory.

prior knowledge of the object, the areal density is not uniquely determined. Instead we use an exponential attenuation approximation ($d \propto ln(\ I/I_0)$) and invert assuming that the object is a single material. We find that density edges are accurately reconstructed using a TV prior, but poorly defined using a smoothing regularizer. Because of the absence of calibration data we do not attempt to recover densities.

3.4 Simulated Example 2

The fourth example is nearly identical to that of the previous example except that MCNP is allowed to correctly account for X-ray scattering. Figure 4 shows (a) a false-color radiograph, (b) computed radial areal density, (c) H^1 regularized reconstruction, and (d) aTV regularized reconstruction. No attempt was made to correct for scattering effects. We find that locations of material boundaries (density discontinuities) are still very well determined. The main effect of the scattering is to render impossible a quantitative analysis of the individual densities. In this relatively noiseless scenario, the material boundaries could be obtained from derivative information on the H^1 solution. However, the previous examples show that this is not generally the case.

3.5 Experimental Example

The fifth example is the reconstruction of a similar spherical comutational test object from an actual X-ray radiograph.[2] Figure 5 shows (a) a false-color radiograph, (b) computed radial areal density, (c) H^1 regularized reconstruction, and (d) aTV regularized reconstruction.. Locations of the actual spherical shell boundaries are indicated by vertical gridlines in the reconstructions. We find

[2]The radiograph was acquired by Hans Snyder and David Miko of the Nuclear Nonproliferation Division at Los Alamos National Laboratory.

that the H^1 solution has a very difficult time distinguishing even qualitative object geometry. The aTV solution is significantly better, identifying all material boundaries. It does, however, retain two intermediate boundaries suggesting a more complex structure than in the real object. We attribute this effect to a combination of the scattering distortion effects, uncertainty in beam intensity correction across the image plane, and measurement noise. The effects of scattering are pronounced in the central region of the reconstructions; the interior and exterior densities of the actual object are both zero.

4 Conclusion

It is expected that any regularization based on prior knowledge of objects to be reconstructed will provide suppression of inversion amplified data noise and visually pleasing results relative to an unregularized solution. The particular choice of regularization, however, has significant influence on the details of the final result. We have shown that TV based regularizations are much better choices for reconstructing objects of piecewise smooth density, especially if the location of density edges are an important reconstruction feature.

We have presented methods for the applying TV regularization to 1D Abel inversion problems with regularization in the 2D object space. Our approach involves the use of a lagged diffusivity fixed point anlaysis. We also introduced a simple adaptive gradient operator approach for identifying density discontinuities and applying zero-cost penalty in the edge location set of the regularization cost functional.

Current and future work includes (1) the use of data fidelity that reflects known data characteristics, (2) extensions to 2D cylindrically-symmetric object descriptions $\rho(r, z)$, (3) arbitrary orientation of the object symmetry axis, and

(4) arbitrary beam geometry (nonparallel).

Acknowledgment

This work was supported by funding from the Department of Energy in the Radiographic Requirements Program and the Technology Integration Program.

References

[1] T. J. Asaki, R. Chartrand, K. R. Vixie, and B. Wohlberg, *Abel inversion using total-variation regularization.* Submitted for publication., 2005.

[2] T. F. Chan and P. Mulet, *On the convergence of the lagged diffusivity fixed point method in total variation image restoration*, SIAM J. Numer. Anal., 36 (1999), pp. 354–367.

[3] A. C. Kak and M. Slaney, *Principles of Computerized Tomography*, Society for Industrial and Applied Mathematics, Philadelphia, 2001.

[4] C. R. Vogel, *Computational methods for inverse problems*, Society for Industrial and Applied Mathematics, Philadelphia, 2002.

Figure Captions

Figure 1. Regularized inversion examples on a synthetic data set generated from a simple nested-ring computational test object. The object density is defined at 200 radial positions and consists of 10 individual density values. The density profile of the object (a) is projected onto 1D and Gaussian noise is added to obtain a synthetic 1D radiograph (b). The remaining subfigures are four regularized object density reconstruction examples (red) against the object reference (black): (c) unregularized; (d) H^1; (e) TV; and (f) aTV.

Figure 2. Regularized inversion examples on a synthetic data set generated from a variable density computational test object. The object density is defined at 200 radial positions. The density profile of the object (a) is projected onto 1D and Gaussian noise is added to obtain a synthetic 1D radiograph (b). The remaining subfigures are four regularized object density reconstruction examples (red) against the object reference (black): (c) unregularized; (d) H^1; (e) TV; and (f) aTV.

Figure 3. Edge detection example from a simulated X-ray experiment. In this example the simulation does not account for X-ray scattering. The four subfigures are: (a) false color radiograph; (b) computed radial areal density; (c) H^1 regularized inversion; and (d) aTV inversion. Locations of actual object density discontinuities are shown as vertical gridlines. The horizonatal gridline indicates the zero density level.

Figure 4. Edge detection example from a simulated X-ray experiment. This example includes the effects of X-ray scattering. The four subfigures are: (a) false color radiograph; (b) computed radial areal density; (c) H^1 regularized

inversion; and (d) aTV inversion. Locations of actual object density discontinuities are shown as vertical gridlines. The horizonatal gridline indicates the zero density level.

Figure 5. Edge detection example on a spherical nondestructive testing example object examined by an actual X-ray experiment. The four subfigures are: (a) false color radiograph; (b) computed radial areal density; (c) H^1 regularized inversion; and (d) aTV inversion. Locations of actual object density discontinuities are shown as vertical gridlines. The horizonatal gridline indicates the zero density level.

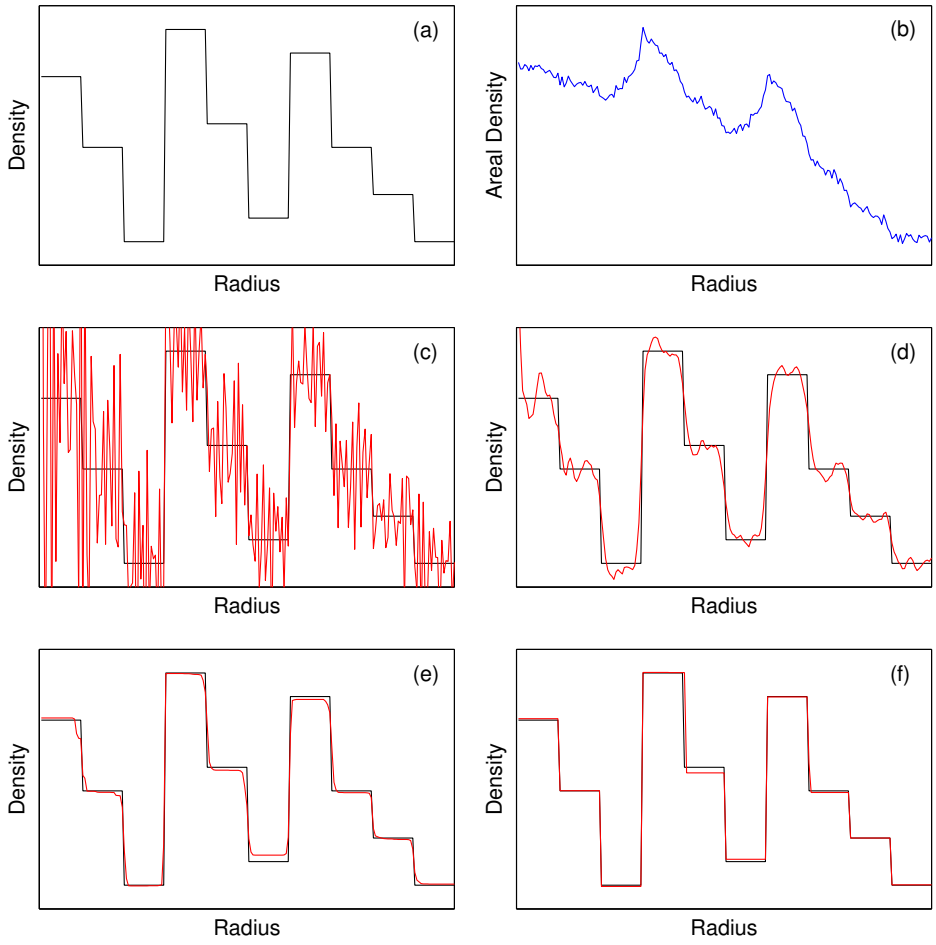
(a)
Density
Radius
(b)
Areal Density
Radius
(c)
Density
Radius
(d)
Density
Radius
(e)
Density
Radius
(f)
Density
Radius

Figure 1:

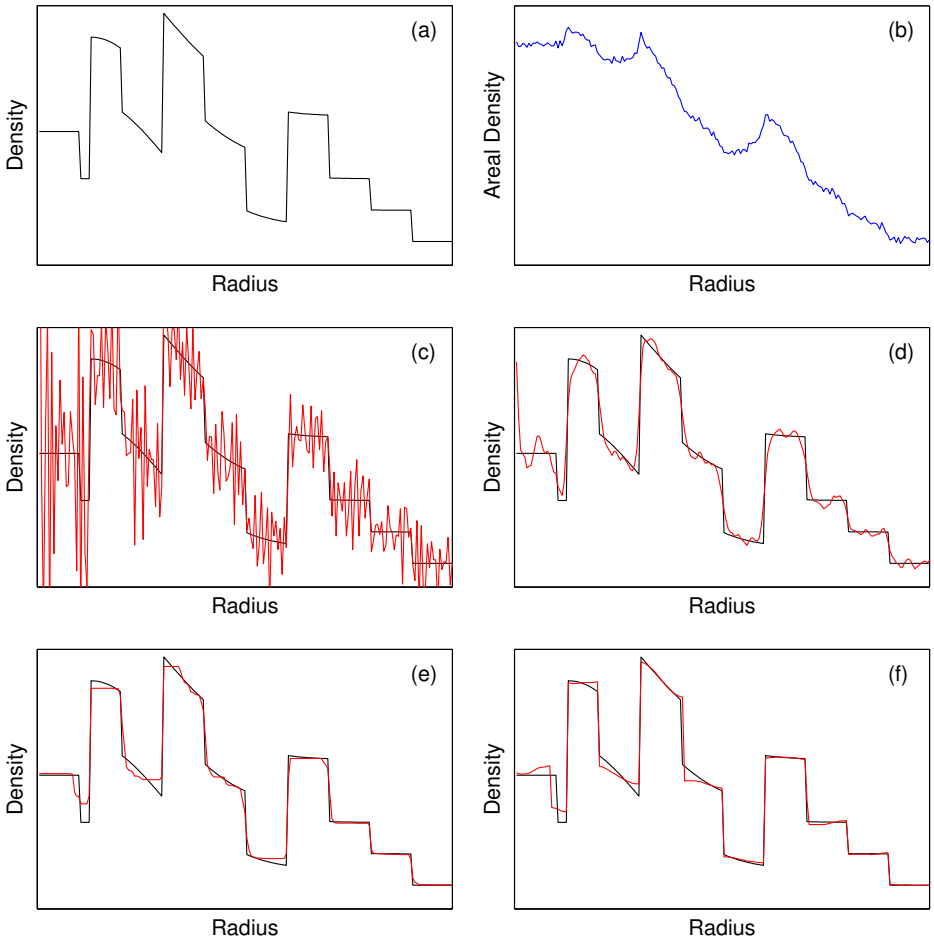
(a)
Density
Radius
(b)
Areal Density
Radius
(c)
Density
Radius
(d)
Density
Radius
(e)
Density
Radius
(f)
Density
Radius

Figure 2:

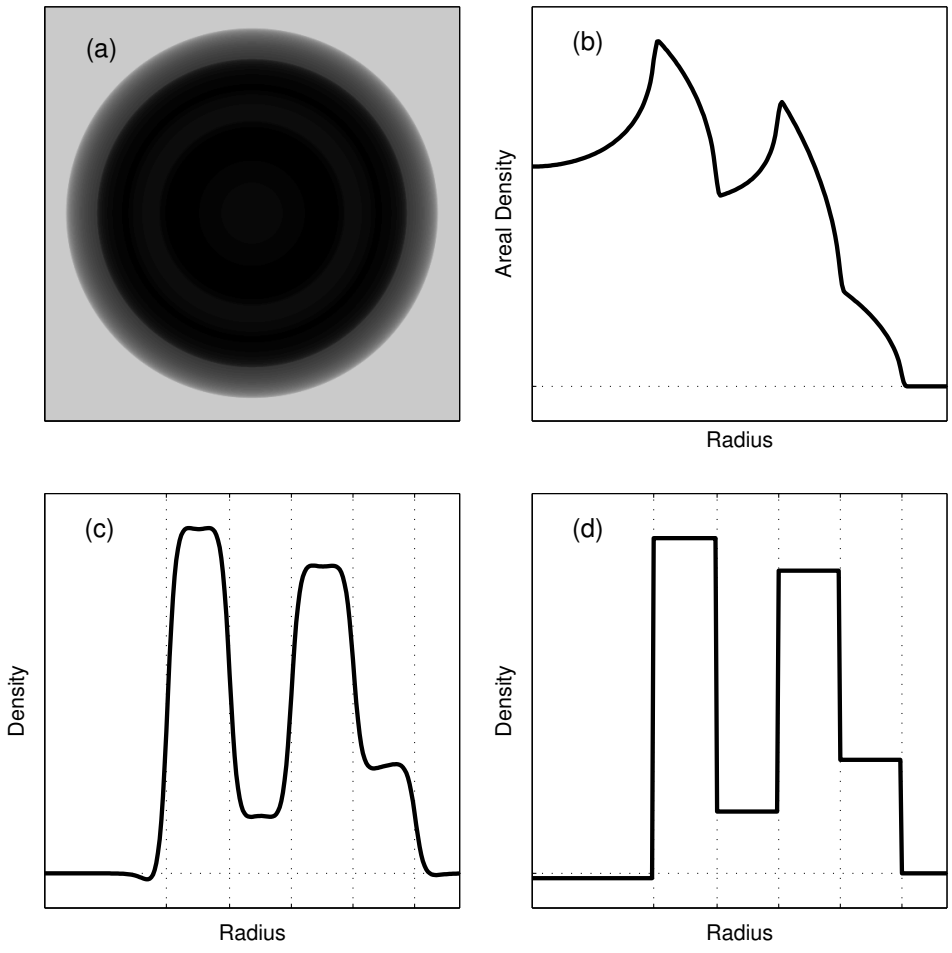

Figure 3:

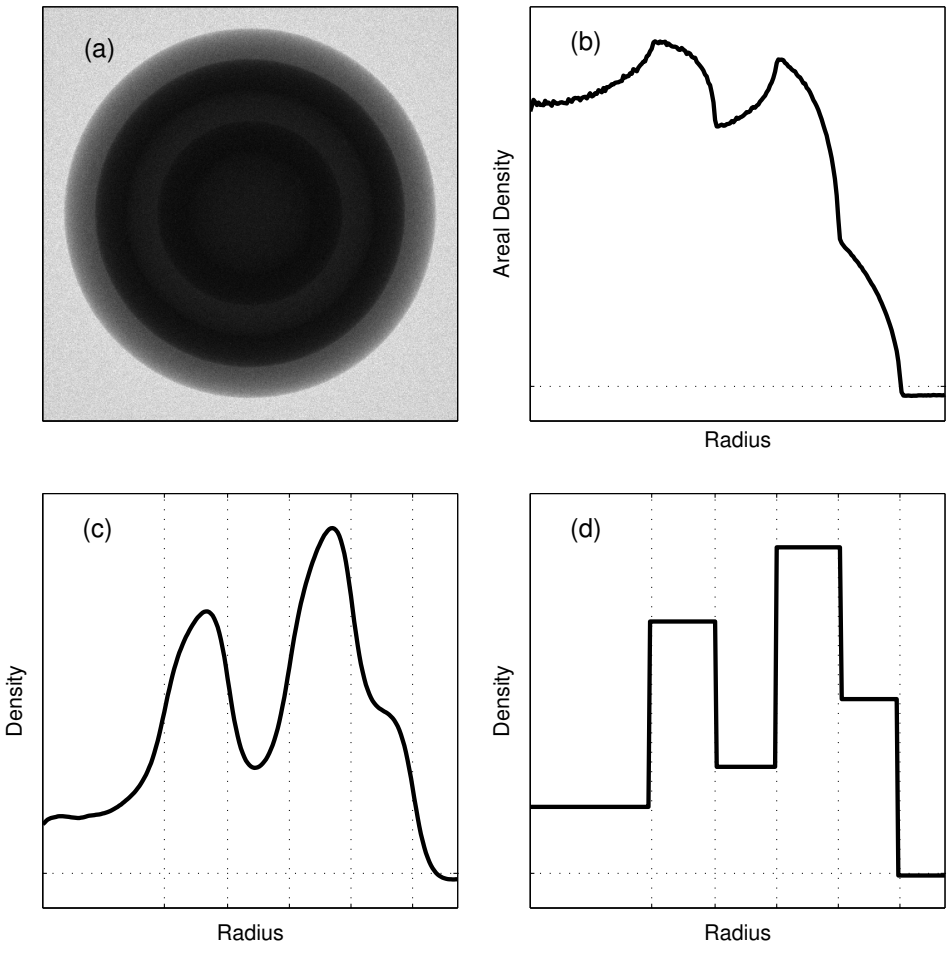

Figure 4:

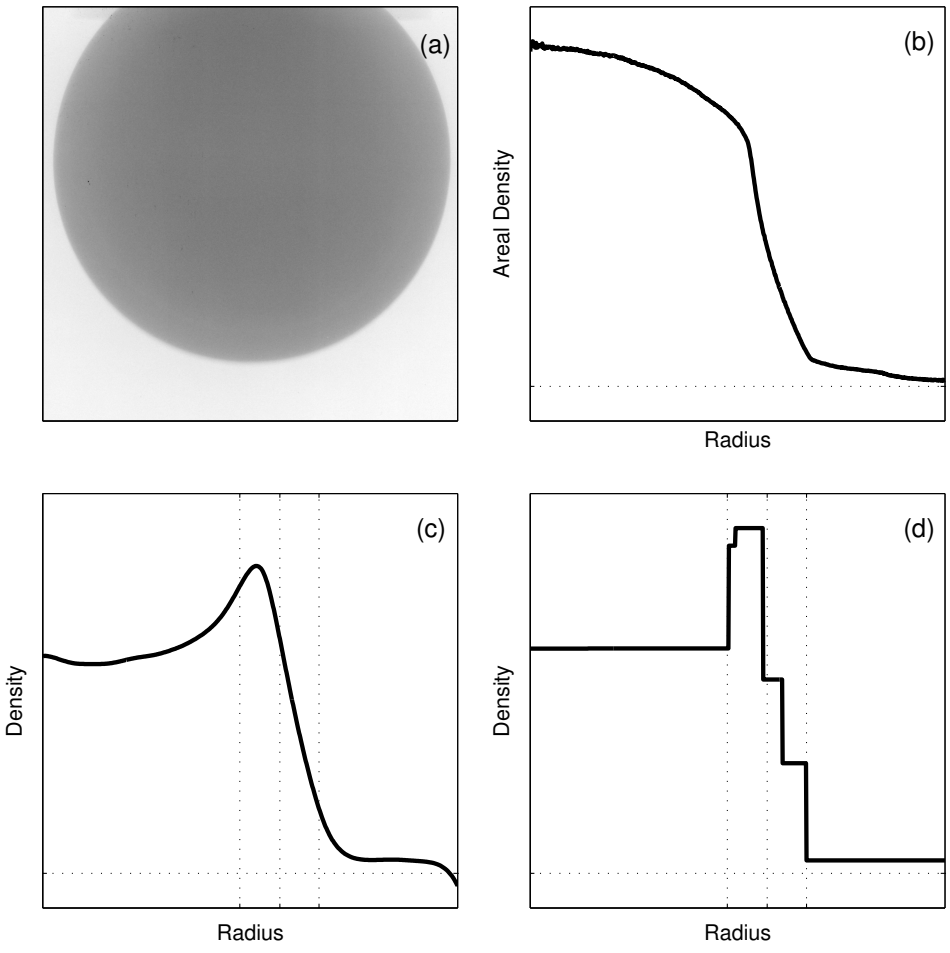

Figure 5: